(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Publication number: **0 373 962 A2**

# (12) EUROPEAN PATENT APPLICATION

(21) Application number: 89313157.3

(22) Date of filing: 15.12.89

(51) Int. Cl.5: C12N 15/52, C12N 9/00, C12N 15/54, C12N 9/12, C12Q 1/527

(30) Priority: 16.12.88 US 285128

(43) Date of publication of application: 20.06.90 Bulletin 90/25

(84) Designated Contracting States: AT BE CH DE ES FR GB GR IT LI LU NL SE

(71) Applicant: BIOTECHNICA INTERNATIONAL, INC.
85 Bolton Street
Cambridge Massachusetts 02140(US)

(72) Inventor: Backman, Keith C.
200 Carlisle Road
Bedford Massachusetts 01730(US)
Inventor: Rudd, Edwin A.
52 Brookdale Road
Salem New Hampshire 03079(US)
Inventor: Lauer, Gail
13 Berkeley Street
Cambridge Massachusetts 02138(US)
Inventor: McKay, Diane
76 Thorndike Street
Arlington Massachusetts 02174(US)

(74) Representative: Deans, Michael John Percy et al
Lloyd Wise, Tregear & CO. Norman House
105-109 Strand
London WC2R OAE(GB)

## (54) Isolating thermostable enzymers.

(57) A thermostable enzyme free from unwanted contaminants is isolated by:

(a) providing a mesophilic host cell engineered to express a gene encoding a heterologous thermostable enzyme;

(b) culturing the mesophilic host cell and producing the thermostable enzyme in a mixture comprising unwanted contaminants; and

(c) purifying the thermostable enzyme by a method that includes heating the enzyme and the unwanted contaminants to a temperature sufficient to inactivate the unwanted contaminants but not sufficient to inactivate the thermostable enzyme.

A method of assaying for nucleic acid ligase is described which includes forming an adenyl adduct with the ligase and detecting the adduct. The latter method is used to screen a library of transformants for the ability to express the nucleic acid ligase.

EP 0 373 962 A2

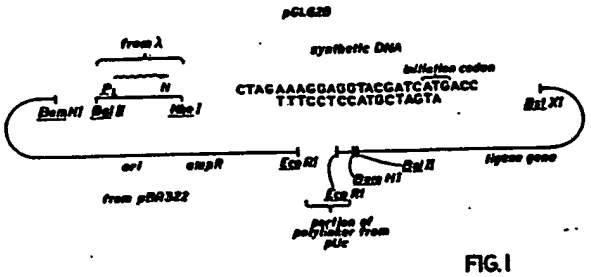


FIG.1

# ISOLATING THERMOSTABLE ENZYMES

This invention relates to the field of producing and recovering thermostable enzymes, particularly enzymes having activity related to nucleic acid metabolism (e.g. RNA or DNA polymerases, ligases, endonucleases or exonucleases).

There are various uses for thermostable enzymes having nucleic acid metabolic activity.

Our European Patent Application No. 88311741.8 (published under No. EP-A-0320308), the disclosure of which is to be regarded as hereby incorporated by reference, discloses an assay in which DNA ligase is used in repetitive cycles of first hybridizing probe pairs to a target, then ligating the hybridized probes, and finally denaturing the hybridized ligated probes. By such cycles, the ligated probes are amplified and detected, so as to provide an amplified signal representative of the target. In this assay, it is desirable to use a thermostable DNA ligase, e. g. from Thermus thermophilus, which can withstand the high temperature required for denaturing the DNA without substantial loss of ligase activity, so that the next ligation cycle is accomplished without adding additional ligase.

U.S. Patents 4,683,195 and 4,683,202 describe a method of amplifying target DNA in a sample, using primers, triphosphates, and DNA polymerase. EP application 0258017 discloses isolation of thermostable DNA polymerase from Thermus aquaticus for use in that amplification procedure. The use of a thermostable enzyme in the procedure is said to avoid the need to add additional enzyme after each denaturation step. EP '017 also reports cloning a Thermus aquaticus DNA polymerase gene and expressing it in E. coli strain DG116 (ATCC 53606).

Takahashi et al. (1984) J. Biol. Chem. 259:10041-10047, describe the purification and properties of the thermostable DNA ligase from Thermus thermophilus. Other references, Romaniec et al. (1987) J. Gen. Microbiol. 133:1297-1308, Liao et al. (1986) Proc. Natl. Acad. Sci. USA 83:576-580, Hara-Yokoyama et al. (1984) J. Biochem. 96:1599-1608, Matsumara et al. (1985) J. Biol. Chem. 260:15298-15303 describe the purification and properties of various thermostable enzymes.

This invention has arisen in the course of our work seeking to provide a straightforward way to separate thermostable enzymes from the contaminants that accompany their in vivo production. When separating the thermostable enzyme from its natural background, troublesome thermostable contaminants are difficult to avoid. For example, when purifying a thermostable ligase or polymerase, it may be difficult to avoid proteases which may be detrimental to the amplification procedures described above. Other thermostable DNA metabolizing enzymes, such as DNA endonucleases also may seriously undermine the effectiveness of the above-described amplification assays. Even if the gene encoding the thermostable enzyme is cloned into a non-thermostable background, the host will produce its own DNA metabolizing enzymes, and proteases. Thus, cloning by itself, only trades one set of contaminants for another.

As will become clear from the detailed description which follows, our techniques provide an opportunity to effectively denature and remove host organism proteins, quickly and easily, possibly in as little as a single step with minimal risk to the desired thermostable enzyme. This enables an abundant, and relatively low-cost source of thermostable enzymes, without detrimental contaminants that otherwise may be difficult to remove.

In accordance with a first aspect of the present invention, there is provided a method for obtaining a thermostable enzyme essentially free from unwanted contaminants, characterised in comprising the steps of:

(a) providing a mesophilic host cell engineered to express a gene encoding a heterologous thermostable enzyme;

(b) culturing said mesophilic host cell to produce said thermostable enzyme in a mixture comprising unwanted contaminants; and

(c) purifying said thermostable enzyme said purification comprising at least one step in which a mixture comprising said unwanted contaminants is heated to a temperature sufficient to inactivate said unwanted contaminants but not sufficient to inactivate said thermostable enzyme.

In this context, purifying the thermostable enzyme means reducing the abundance or activity level of contaminating substances (e.g. endonucleases, exonucleases, proteases) particularly those which interfere with the intended use of the enzyme such as in the selective amplification procedure described above, or with recovery of the enzyme. Mesophilic host cells are cells which can be engineered to produce the desired thermophilic enzyme and whose proteins generally are denatured at a temperature that does not denature the desired thermophilic enzyme. Representative suitable mesophilic host cells are described below. Heterologous, as used above, means that the enzyme is not naturally produced by the host cell.

In preferred embodiments, the mesophilic host cells are lysed after culturing, to produce a lysate comprising the thermostable enzyme and unwan-

ted contaminating proteins. Preferably, the heating step essentially completely denatures and precipitates the undesired contaminants, so that the contaminants are readily separated from liquid phase containing the desired thermostable enzyme. The method preferably is used to purify enzymes involved in metabolism of, modification of, or action on a nucleic acid, for example nucleases, nucleic acid methylases, polymerases, ligases, or recombinases. Particularly preferred enzymes are DNA ligases or polymerases. Preferred enzymes are those from thermophilic bacteria (described in more detail below) which survive a heating step of 65° C to about 100° C (most preferably 80° C-95° C), for a sufficient period of time (e.g. at least about 1-3 minutes, and preferably for at least 5 minutes) to inactivate mesophilic host cell proteins.

Our described techniques overcome cloning-related difficulties experienced in obtaining thermostable ligases, such as the difficulty of easily assaying for ligase activity when screening transformants, and the difficulty of obtaining high levels of ligase expression in transformants.

In particular, a second aspect of the invention features a method of assaying for NAD-dependent nucleic acid ligase activity (particularly thermostable ligase activity) by forming a stable covalent adenyl (e.g from labeled NAD) adduct with the ligase, and detecting the adduct. Ligase activity present in a sample comprising undesired contaminants can be detected by heating the sample under conditions which denature the contaminants. This is particularly useful where any contaminants which would also form adenyl adducts are denatured.

This assay method may be used to screen host cells transformed with a library of nucleic acid segments, thereby identifying cells which express ligase activity.

A third aspect of the invention generally features a cloned thermostable nucleic acid ligase gene, e.g. nucleic acid encoding a protein having nucleic acid (DNA) ligase activity which is thermostable, wherein the nucleic acid is part of an engineered cloning vehicle in which that gene does not naturally occur.

A fourth aspect of the invention features an essentially pure DNA sequence encoding a thermostable nucleic acid ligase.

Preferably, in the third and fourth aspects, the cloned gene is part of an expression vehicle (a plasmid or other transformable DNA segment) comprising a promoter that provides enhanced (e.g. at least twice the level of gene product in unengineered cells) expression of the gene.

Other features and advantages will be apparent from the following description.

Fig. 1 is a diagrammatic representation of the components of plasmid pGL628 that contains a Thermus thermophilus DNA fragment extending from a BstXI site to a BglII site about 2.6 kb away.

Those skilled in the field will recognize that there are numerous ways to perform the various steps in the method for a wide variety of enzymes. Generally, the techniques of recombinant DNA manipulation, expression and cell culturing are well-known techniques understood in the art. The mesophilic host cells used in the method can be procaryotes represented by various strains of E. coli. However, other microbial strains may also be used, such as bacilli, for example Bacillus subtilis, various species of Pseudomonas, or other bacterial strains. In such procaryotic systems, plasmid vectors that contain replication sites and control sequences derived from the host or from a species compatible with the host are used. In addition to bacteria, eucaryotic microbes, such as yeast, may also be used as hosts. Laboratory strains of Saccharomyces cerevisiae, Bakers' yeast, are most often used, although a number of other strains are commonly available

It is also of course, possible to express genes encoding the thermophilic enzymes in eucaryotic host cell cultures derived from multicellular organisms. See, for example, Tissue Culture, Academic Press, Cruz and Patterson, editors (1973). Useful host cell lines include murine myelomas N51, VERO, MD canine kidney cells, HeLa cells, and Chinese hamster ovary (CHO) cells. Expression vectors for such cells ordinarily include promoters and control sequences compatible with mammalian cells such as, for example, the commonly used early and late promoters from Simian Virus 40 (SV 40) (Fiers, et al., Nature (1978) 273:113), or other viral promoters such as those derived from polyoma, Adenovirus 2, bovine papilloma virus, or avian sarcoma viruses, or immunoglobulin promoters and heat shock promoters. A system for expressing DNA in mammalian systems using the BPV as a vector is disclosed in U S. Patent 4,419,446. A modification of this system is described in U.S. Patent 4,601,978. General aspects of mammalian cell host system transformations have been described by Axel, U.S. Patent No. 4,399,216. Origins of replication may be obtained, if needed, from viral sources. However, integration into the chromosome is a common mechanism for DNA replication in eucaryotes.

Plant cells are also now available as hosts, and control sequences compatible with plant cells such as the nopaline synthase promoter and polyadenylation signal sequences (Depicker, A., et al., J.Mol. Appl. Gen. (1982) 1:561) are available.

Recently, in addition, expression systems employing insect cells utilizing the control systems provided by baculovirus vectors have been de-

scribed (Miller, D.W., et al., in Genetic Engineering (1986) Setlow, J.K. et al., eds., Plenum Publishing, Vol. 8, pp. 277-297).

Construction of suitable vectors containing the desired coding and control sequences employs standard ligation and restriction techniques that are well understood in the art. Isolated plasmids, DNA sequences, or synthesized oligonucleotides are cleaved, tailored, and religated in the form desired.

Site-specific DNA cleavage is performed by treating with the suitable restriction enzyme (or enzymes) under conditions that are generally understood in the art, and the particulars of which are specified by the manufacturer of these commercially available restriction enzynes. See, e.g., New England Biolabs, Product Catalog.

Cloning of genes encoding thermostable nucleic acid metabolic genes, e.g. T. thermophilius DNA ligase, is significantly assisted by development of an effective method for detecting the activity of such enzymes. In particular, where the enzymes are expressed only at extremely low levels in the cloning background, detection may be so difficult that traditional strategies predicated on identifying clones which express the desired gene (e.g. complementation of deficient mutants or immunological identification of desired transformants) are unlikely to be effective without such an improved screening method.

The problem is exacerbated by lack of characterization of the desired protein, which hampers cloning strategies based on sequence data. In particular, T. thermophilus ligase has not been sequenced because of poor yields of enzyme and impurities (particularly proteolytic impurities).

Accordingly, Example 1 regarding a method for detecting the nucleic acid metabolic enzyme (particularly thermostable DNA ligase) is an important aspect of our cloning strategy. Other Examples are provided concerning cloning and expression. These Examples are provided to illustrate the invention, but not to limit it.

## Example 1 - Detection of the Thermophilic Enzyme

The following observations provide a method of detecting a thermophilic enzyme which can be labeled with NAD, e.g. Thermus thermophilus ligase, in a mixture of host proteins.

NAD is most often utilized as a mediator of oxidation/reduction reactions, and very few enzymes utilize NAD as a source of stored chemical energy Prokaryotic DNA ligases are one class of enzyme which does. Most other energy consuming enzymatic processes, including certain DNA ligases, utilize ATP as an energy source. Very few prokaryotic enzymes survive and remain active at temperatures in excess of 70°C. Aside from idiosyncratic examples, the majority of such enzymes exist in thermophilic organisms.

It is known that some nucleic acid metabolic enzymes (e.g. DNA ligase) form a covalently linked enzyme adenylate species as an intermediate in the reaction sequence they catalyze. We have discovered that at temperatures between 70°C and 80°C, thermostable DNA ligase from the thermophile Thermus thermophilus can form a stable, covalent adenyl adduct, and that the adduct can be specifically labelled with $^{32}P$ if the NAD used as a cofactor is labelled with $^{32}P$ (available from New England Nuclear). The enzyme has an extremely low $K_M$ for NAD, so the above reaction is satisfactory for labelling the cloned ligase product, even though extremely small quantities of thermophilic ligase are expressed in E. coli prior to engineering for increased expression.

In general, mesophilic host proteins either cannot be labeled at all with NAD, or their ability to be labeled is substantially eliminated by previous exposure to 70°C or higher

In outline, extracts suspected of containing thermophilic ligase are incubated at 70°C or higher; $^{32}P$-NAD is added; samples are incubated further; samples are denatured and run on SDS polyacrylamide gel electrophoresis; the gel is autoradiographed; and the presence or absence of DNA ligase in a sample is ascertained by the presence or absence of a labelled species which comigrates with authentic ligase.

More specifically, cells are grown and cell extracts are assayed as follows.

The cells from a plate containing ≤ 70 colonies transformed with a library of cloned DNA segments from T. thermophilus are scraped into 3.0 ml of L broth. This suspension was used to inoculate 100 ml of L broth in a 2 liter flask. The cells were grown at 37°C until they had reached saturation (~ 3-4 hrs.). The cells from a 20 ml aliquot of the culture were harvested by centrifugation, and the cell pellet was stored at -20°C.

The cell pellets described above were suspended in 10 ml of disruption buffer (20mM TRIS, 0.1mM EDTA and 1mM β-mercaptoethanol, pH 7.6). The cells were disrupted by sonication. The cell debris was removed by centrifugation (20 min. at 20,000 rpm, 4°C, SS-34 rotor). The resulting supernatant was incubated in a 80°C water bath for 30 minutes. The precipitated proteins were then removed by centrifugation (20 min., 7,000 rpm at 4°C, SS-34 rotor). The resulting supernatant was placed in an Amicon Centriprep™ (10,000 MWCO) and concentrated to ≈ 0.5 ml. The concentrate was transferred to an Amicon Centricon™ (10,000 MWCO) and the process of concentration was continued until a volume of ≈ 50 µl was reached. The

concentrated samples were then placed in plastic microfuge tubes and incubated in an 80°C water bath for 20 minutes. Precipitate was again removed by centrifugation (4 min. in microfuge) and the supernatant was transferred to another plastic microfuge tube. The tube was again placed in the 80°C water bath After the sample had reached temperature, 1.0 µl of diluted $^{32}$P-NAD was added and the tube was incubated for 2 hours at 80°C (NEN stock $^{32}$P-NAD was diluted 1/2 in 0.2M TRIS, pH 7.6). Then 25 µl of Laemmli SDS gel sample buffer was added to the tube and the incubation at 80°C was continued for another 20 minutes. The sample was then electrophoresed on a 10% Laemmli SDS-gel. When the tracking dye neared the bottom of the gel the electrophoresis was stopped, and the bottom of the gel containing most of the $^{32}$P label (NAD) was cut away. The gel was soaked in $H_2O$ for 30 minutes and then dried on a gel dryer. The dried gel was then autoradiographed. The antoradiograph was analyzed for the presence of species which comigrated with authentic ligase during electrophoresis.

## Example 2 - Assay of DNA Ligase

Ligase can be semi-quantitatively assayed by a nick sealing procedure. Plasmid DNA (Puc19) is nicked by treatment with BamHI endonuclease in the presence of ethidium bromide. (pUC18 and many other plasmids could also be used.) Ligase is assayed by treatment of a fixed amount of nicked plasmid with dilutions of samples containing ligase, followed by agarose gel electrophoresis in the presence of ethidium bromide. Conversion of nicked to CCC form DNA is assessed and correlated with dilution. The assay is especially useful for determining thermophilic ligase activity when the reaction is run at a temperature between 50°C and 75°C.

The specific DNA ligase assay is conducted in the following buffer: 50mM Tris, 10mM $MgCl_2$, 50mM KCl, 2mM $MnCl_2$, 1mM DTT, 10µg/ml BSA, and 0.1mM NAD. Ligation is performed using a mixture that is 2µl DNA ligase containing sample, 1µl DNA (1µg/ml nicked DNA), and 7µl of the above buffer. The mixture is incubated for 30 min. at 68°C and ligation is stopped using 6µl of 50mM EDTA/0.5% SDS, and 0.05% BPB.

To detect ligase, 5µl of the 16µl mix is electrophoresed on a gel of 0.8% agarose with 1.0µl/ml ethidium bromide, TBE running buffer. A positive control is run using E. coli DNA ligase

## Example 3 - Cloning of Thermophilic Ligase

T. thermophilus cells (ATCC 27634) were grown and DNA was prepared from gently lysed cells by equilibrium density gradient centrifugation in CsCl. DNA was partially digested with Sau3AI and cloned in the BclI site of the positive selection vector pTR264, a version of pTR262 (Roberts et al., Gene 12:123 (1980)) which contains an ampicillin resistance gene. Tetracycline resistant clones (i.e. those having inserts) were isolated. Pools of 70 colonies were made, and lysates of such pools were examined for the presence of thermophilic ligase by the procedure of Example 1. A positive pool was identified, and broken into sub-pools of 7 colonies. By repeating the procedure and breaking down the subpool to single colonies, a clone containing thermophilic ligase was identified. It carried a plasmid which was named pGL600

The identity of the NAD-labelling species was confirmed to be DNA ligase by assaying as described in Example 2.

## Example 4 - Location of DNA Ligase Gene

PGL600 contained about 8kb of inserted DNA. By subcloning, we localized the gene to a region of about 4kb (the coding capacity required to specify the gene was estimated to bo about 2.2 kb). Since the subcloning was done in pUC18 and PUC19, identical subclones were obtained which were positioned oppositely with respect to an external lac promoter. Quantitation of the levels of ligase produced by such pairs of clones strongly suggested the orientation of the ligase gene.

Deletion of a BglI bounded fragment near one end eliminated ligase formation and suggested that the beginning of the gene lay within that fragment.

The BglII generated fragment was sequenced, and a putative start point for the ligase gene was identifed by analysis of the sequence.

## Example 5 - Expression of Thermophilic Ligase in E. coli

Although thermophilic ligase was detectably expressed in E. coli, the low and variable levels suggested to us that considerable improvement could be obtained. Analysis of the sequence revealed three useful facts.

1. The putative start codon is cleaved by the restriction endonuclease BstXI (See Fig. 1); further analysis revealed that this site was unique within the DNA spanning the gene;

2. The gene lacked an optimal ribosome binding site for expression in E. coli; and,

3. The gene apparently lacked an associated promoter conforming to the consensus promoter

sequence in E. coli.

An appreciation of these observations informed the following procedure:

A DNA fragment comprising almost all of the gene was prepared from the BstXI site at one end to a BglII site past the distal end of the gene. Synthetic DNA was prepared which could link the prepared DNA to a strong promoter; this synthetic DNA restored the initiator codon and provided a good ribosome binding site (See Fig. 1).

Several such constructions were made, varying the nature of the promoter and the ribosome binding site. For example the lac promoter from the pUC plasmid was used, the λ $P_L$ promoter was used, and the M13 gene II promoter was used. DNA ligase production was monitored by polyacrylamide gel electrophoresis of cell lysates and by the assay procedure of Example 2.

One suitable construction resulting from the above procedure is a plasmid designated pGL628. Its components are shown in Fig. 1.

Construction of pGL628 was accomplished in cells which provided cI repressor to control expression from the $P_L$ promoter on the plasmid U.S. patent application SN 091,837, filed September 1, 1987 by Keith C. Backman and Barbara Bolten, describes a method for preparing such strains. One such strain is KB649.

Plasmid pGL628 contains the ~2.6 kb BstXI to Bgl II Thermus thermophilus DNA ligase DNA fragment. That plasmid, in E. coli, strain KB649 has been deposited in the American Type Culture Collection, Rockville, MD and given ATCC Accession No. 67858.

## Example 6 - Production of Ligase

Strain KB649/pGL628 was grown in a fermentor at 30°C in 10 litres of minimal salts (e.g. M9 media) plus 15-20g/l glucose. When the cells reached mid log phase, the temperature set point was changed to 42°C, and incubation was continued for about 2 hours. Cells were harvested, and a cell paste was stored frozen. By examining cell lysates resolved by polyacrylamide gel elecrophoresis in the presence of sodium dodecyl sulfate, it was determined that between 5% and 10% of the total cell protein was thermophilic ligase.

## Example 7 - Ligase Purification from pGL628

Cell pellets (33g) were resuspended at 4ml/g in disruption buffer. Disruption buffer consisted of the following: 50mM Tris-HCl, 10mM $MgCl_2$, 50mM KCl, 2mM $MnCl_2$, 1mM DTT and 10mM $NH_4Cl$, pH7.7

The disruption was carried out in a French Pressure cell at 14,000 PSI (9.65 x $10^7$ N/m²).The suspension was then spun down at 20K rpm for 20 minutes. The supernatant was removed and incubated in an 80° C $H_2O$ bath for 30 minutes. This suspension was then spun at 7.5K rpm for 15 minutes. The supernatant was then made up to 2% streptomycin sulfate using a 10% stock solution. This 2% solution was stirred on ice for 20 minutes, then spun at 12.5K rpm for 20 minutes.

This supernatant was loaded on a Cibacron blue column (136ml bed volume) equilibrated with 20mM Tris, 50 mM KCl, pH7.6. The column was eluted with a three-column volume gradient of 50 to 500 mM KCl. Fractions were collected in a 5 ml volume at a rate of 2.0 ml/min.

Active fractions were pooled, and subjected to an Amicon stirred cell concentrator. then an Amicon centriprep concentrator until the volume was 1-5% of the following gel filtration column. Specifically the sample was loaded on a Fractogel 50S column (bed volume 100 ml) equilibrated with 10mM $KPO_4$, 100mM KCl, 1mM $MgCl_2$, pH7.6, 5ml fractions were collected, at a rate of 0.5ml/min.

Active fractions were pooled, concentrated in an Amicon Centriprep, made to 20% glycerol and stored at -20° C.

Other embodiments are feasible.

For example, other thermostable enzynes such as polymerases can be prepared.

## Claims

1. A method for obtaining a thermostable enzyme essentially free from unwanted contaminants, characterised in comprising the steps of:

(a) providing a mesophilic host cell engineered to express a gene encoding a heterologous thermostable enzyme;

(b) culturing said mesophilic host cell to produce said thermostable enzyme in a mixture comprising unwanted contaminants; and

(c) purifying said thermostable enzyme, said purification comprising at least one step in which a mixture comprising said unwanted contaminants is heated to a temperature sufficient to inactivate said unwanted contaminants but not sufficient to inactivate said thermostable enzyme.

2. A method according to Claim 1, further characterised in that after said culturing step, the mesophilic host cell is lysed to produce a lysate comprising said thermostable enzyme.

3. A method according to Claims 1 or 2, further characterised in that said purification comprises heating said mixture comprising said unwanted contaminants to effect essentially completely pre-

cipitation of said unwanted contaminants from a liquid phase comprising the thermostable enzyme.

4. A method according to Claim 3, further characterised in that said unwanted contaminants comprise proteins, and said heating step denatures said proteins causing them to precipitate.

5. A method according to any preceding claim, further characterised in that said unwanted contaminants are endonucleases, exonucleases, or proteases.

6. A method according to any preceding claim, further characterised in that the thermostable enzyme catalyses metabolism of, modification of, or action on a nucleic acid.

7. A method according to Claim 6, further characterised in that the thermostable enzyme is a DNA ligase or a DNA polymerase.

8. A method according to Claim 7, further characterised in that the thermostable enzyme is a DNA polymerase or a DNA ligase from a thermophilic bacteria.

9. A method according to Claim 8, further characterised in that the thermostable enzyme is from a thermophilic bacterium selected from Thermus flarus, Thermus ruber, Thermus thermophilus, Bacillus stearothermophilus, Thermus aquaticus, Thermus lacteus, Thermus rubens and Methanothermus fervidus.

10. A method according to any preceding claim, further characterised in that said temperature is from 65° C to about 100° C, and preferably from 80° C to about 95° C.

11. A method for assaying for NAD-dependent thermostable nucleic acid ligase activity, characterised in comprising forming a stable covalent AMP adduct of said ligase and detecting said adduct.

12. A method according to Claim 11, further characterised in that the adenyl moiety of said adduct is labeled.

13. A method according to Claims 11 or 12, further characterised in that the adenyl moiety of said adduct is derived from NAD.

14. A method according to any of Claims 11, 12 or 13, further characterised in that said ligase is a thermophilic ligase, and said ligase activity is detected in a sample comprising undesired contaminants: and in that said method comprises heating a mixture comprising said ligase and said undesired contaminants to a temperature which denatures or inactivates said undesired contaminants but not said ligase.

15. A method of cloning nucleic acid encoding nucleic acid ligase activity, in which a library of nucleic acid segments is transformed into host cells, characterised in that said transformed host cells are screened to identify cells which express ligase activity by a method according to any of Claims 11 to 14.

16. A cloned thermostable nucleic acid ligase gene.

17. A gene according to Claim 16, further characterised in that said gene is part of an expression vehicle comprising a promoter providing enhanced expression of said gene.

18. A gene according to Claim 17, further characterised in that said expression vehicle is pGL628.

19. An essentially pure nucleic acid sequence encoding a thermostable nucleic acid ligase.

EP 0 373 962 A2



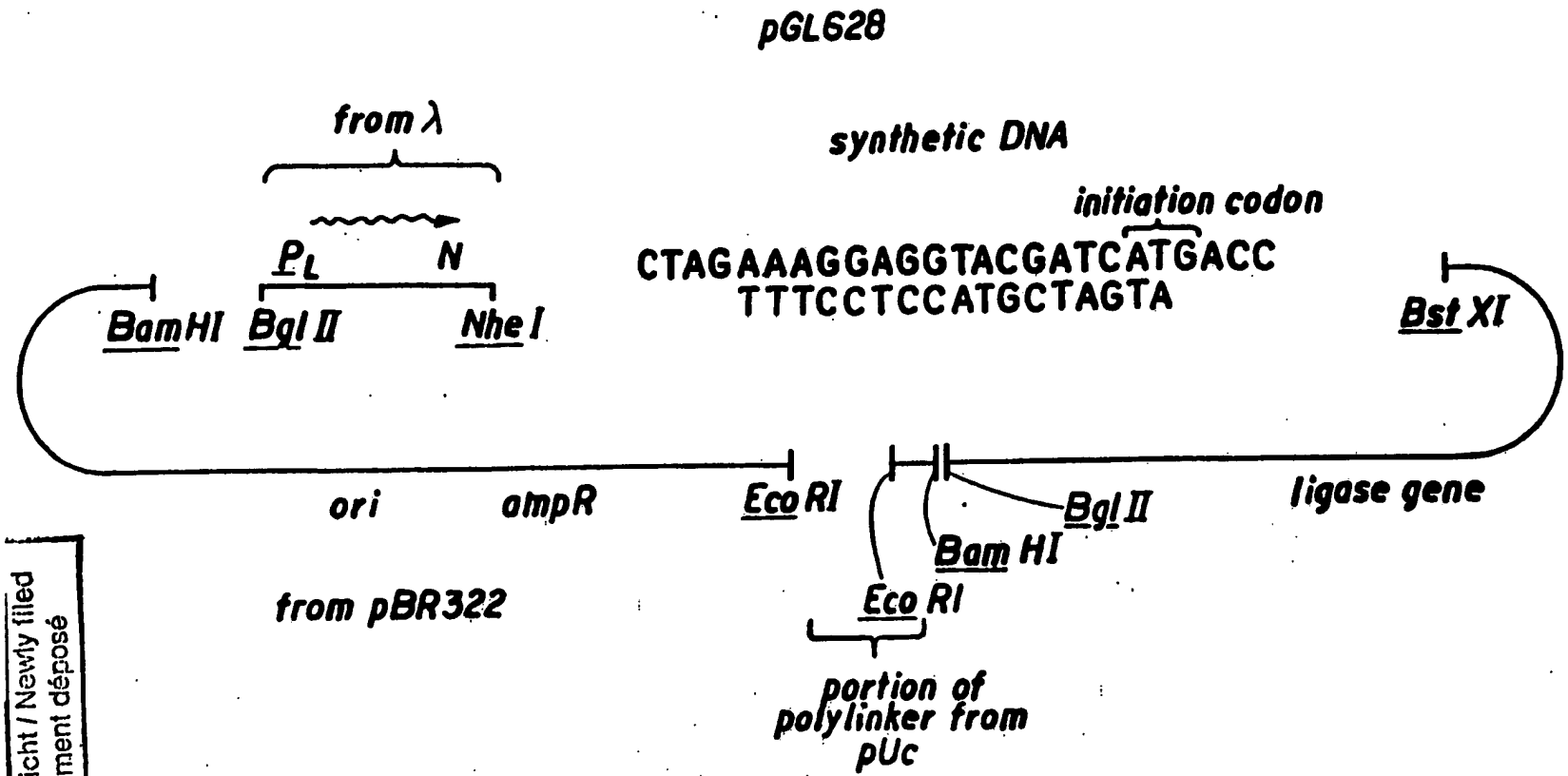


FIG.1

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Publication number: **0 373 962 A3**

# (12) EUROPEAN PATENT APPLICATION

(21) Application number: 89313157.3

(22) Date of filing: 15.12.89

(51) Int. Cl.5: C12N 15/52, C12N 9/00, C12N 15/54, C12N 9/12, C12Q 1/527

(30) Priority: 16.12.88 US 285128

(43) Date of publication of application: 20.06.90 Bulletin 90/25

(84) Designated Contracting States: AT BE CH DE ES FR GB GR IT LI LU NL SE

(88) Date of deferred publication of the search report: 21.08.91 Bulletin 91/34

(71) Applicant: BIOTECHNICA INTERNATIONAL, INC.
85 Bolton Street
Cambridge Massachusetts 02140(US)

(72) Inventor: Backman, Keith C.
200 Carlisle Road
Bedford Massachusetts 01730(US)
Inventor: Rudd, Edwin A.
52 Brookdale Road
Salem New Hampshire 03079(US)
Inventor: Lauer, Gail
13 Berkeley Street
Cambridge Massachusetts 02138(US)
Inventor: McKay, Diane
76 Thorndike Street
Arlington Massachusetts 02174(US)

(74) Representative: Deans, Michael John Percy et al
Lloyd Wise, Tregear & CO. Norman House
105-109 Strand
London WC2R OAE(GB)

## (54) Isolating thermostable enzymers.

(57) A thermostable enzyme free from unwanted contaminants is isolated by:

(a) providing a mesophilic host cell engineered to express a gene encoding a heterologous thermostable enzyme;

(b) culturing the mesophilic host cell and producing the thermostable enzyme in a mixture comprising unwanted contaminants; and

(c) purifying the thermostable enzyme by a method that includes heating the enzyme and the unwanted contaminants to a temperature sufficient to inactivate the unwanted contaminants but not sufficient to inactivate the thermostable enzyme.

A method of assaying for nucleic acid ligase is described which includes forming an adenyl adduct with the ligase and detecting the adduct. The latter method is used to screen a library of transformants for the ability to express the nucleic acid ligase.

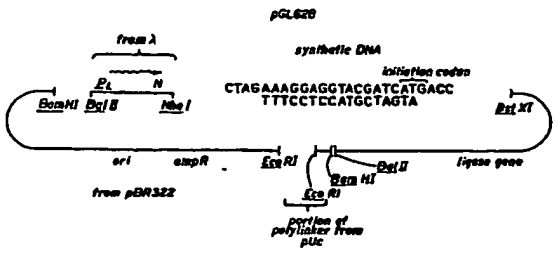


FIG.1

EP 0 373 962 A3

European Patent Office

# EUROPEAN SEARCH REPORT

Application Number

EP 89 31 3157

## DOCUMENTS CONSIDERED TO BE RELEVANT

| Category | Citation of document with indication, where appropriate, of relevant passages | Relevant to claim | CLASSIFICATION OF THE APPLICATION (Int. Cl.5) |
|---|---|---|---|
| D,A | EP-A-258017 (CETUS CORPORATION)<br>* the whole document * | 1-10 | C12N15/52<br>C12N9/00 |
| D,Y | | 16-19 | C12N15/54<br>C12N9/12 |
| Y | CHEMICAL ABSTRACTS, vol. 94, no. 5, 02 February 1981<br>Columbus, Ohio, USA<br>YANG, Q.-S. et al.: "Expression of cloned DNA ligase gene (G30) of E.coli bacteriophage T4 in vivo"<br>page 267; left-hand column; ref. no. 94:27231D<br>* abstract * | 16-19 | C12Q1/527 |
| P,X | WO-A-8906691 (CETUS CORPORATION)<br>* page 69, line 11 - page 71, line 9; claims 25-26 * | 1-10 | |
| D,A | JOURNAL OF BIOLOGICAL CHEMISTRY.<br>vol. 259, no. 16, 25 August 1984, BALTIMORE US<br>TAKAHASHI, M. et al.:<br>"Thermophilic DNA ligase : purification and properties of the enzyme from Thermus thermophilus HB8"<br>pages 10041-10047<br>* the whole document * | 1 | |

TECHNICAL FIELDS SEARCHED (Int. Cl.5)

C12N
C12Q

The present search report has been drawn up for all claims

| Place of search | Date of completion of the search | Examiner |
|---|---|---|
| THE HAGUE | 26 JUNE 1991 | ANDRES S.M. |

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P0401)

Consommation et Corporations Canada
Consumer and Corporate Affairs Canada

Bureau des brevets
Patent Office

Ottawa, Canada
K1A 0C9

(21) (A1) 2,049,879
(22) 1991/08/26
(43) 1992/03/01

5,039,755

(51) INTL.CL.[5] C12N-015/10; C12Q-001/68

(19) (CA) **APPLICATION FOR CANADIAN PATENT** (12)

# (54) Controlled Initial Target-Dependent Production of Templates for Ligase Chain Reaction

(72) Backman, Keith C. - U.S.A. ;
Shimer, George H. - U.S.A. ;
Carrino, John J. - U.S.A. ;

(73) Abbott Laboratories - U.S.A. ;

(30) (US) 575,177 1990/08/30

(57) 8 Claims



Canada

Abstract of the Disclosure

An improvement to the ligase chain reaction (LCR) method of selectively and geometrically ligating nucleic acid probes; before the LCR, a template-dependent controlled reaction increases the population of templates capable of supporting template-dependent ligation in LCR selectively in the presence of target nucleic acid, while substantially avoiding creation of any molecules capable of supporting template-dependent ligation in the absence of target nucleic acid. One way to control the initial reaction is to weight the reaction to create templates corresponding to only one of the two target strands, using e.g. template dependent ligation or polymerization. Another way to control the initial amplification is to create templates corresponding to both strands, but to prevent spurious creation of ligation templates by using a probe concentration low enough to avoid blunt end ligation or by using a method of amplification such as PCR that does not generate spurious ligation templates.

# CONTROLLED INITIAL TARGET-DEPENDENT PRODUCTION OF TEMPLATES FOR LIGASE CHAIN REACTION

This invention pertains to the detection of nucleic acid sequences; specifically, it pertains to target-dependent enhancement of signal relative to background in the detection of nucleic acid sequences by the Ligase Chain Reaction (LCR) technique described below.

An aspect of LCR involves a method for detecting the presence of specific nucleic acid sequences in samples. LCR features target-dependent ligation using two sets of nucleic acid probes. The members of the first set of probes are designed to hybridize to a target strand at abutting locations, in end to end fashion. The abutting probes thus aligned are ligated by template-dependent creation of a phosphodiester bond. The newly created molecule serves as a template (as does target complement if present) to mediate ligation of the second pair of probes to generate a new template for ligation of additional first pairs of probes, and so on.

As described below in more detail, the power of LCR derives from the ability of each set of ligated probes to act as a template for further ligation, thus providing ligation at an increasing (geometric) rate--i.e., an increasing number of molecules containing ligated target sequences is added by each cycle. LCR has been disclosed in EP-A-320 308. Because LCR generates a specific phosphodiester linkage and thereby amplifies the number of templates for creating such a linkage in further cycles, we may sometimes characterize LCR as an amplification technique. We do not mean to say that LCR amounts to amplification in the sense of synthesis of nucleic acid. LCR products have the sequence information of the probes provided, which is not necessarily identical to that of the target.

Ideally, the above-described template-dependent ligation occurs if and only if the template is present. In

fact, however, template-independent ligation may occur as well. Template-independent ligation events create amplifiable products that can result in a background signal even in the absence of target. More specifically, LCR probe sets generally are in significant excess of the target DNA throughout most of the process. These excess complementary probes are free to hybridize to one another creating blunt end duplex DNA molecules which can be joined together by DNA ligase independently of target (albeit at substantially lower efficiencies than target mediated ligation), resulting in the generation of spurious templates that can then enable a chain reaction of template-dependent ligation and lead to a spurious "background" signal in LCR.

One method for controlling spurious background in LCR amplification involves modifying the ends of the probes to prevent ligation (including undesired blunt end ligation), and then reversing the modification in a template-dependent fashion. See EP-A-0 439 182 by Backman, Bond, Carrino and Laffler.

Wu, et al., in Polymerase Chain Reaction, Erlich, et al., Eds, (Cold Spring Harbor Laboratory Press, 1989), p. 233-236, describe allele specific amplification schemes. At one point, they suggest that LAR ("Ligase Amplification Reaction"--another name for LCR) may be used as an allele specific detection system for PCR enriched DNA sequences. It appears they propose to use PCR as an initial template controlled reaction, prior to LCR.

Wu, et al., Genomics, 4:560-569 (1989) review various amplification methods. They discuss both LAR (LCR) and PCR. In one variation, they describe a linear LCR reaction, employing just one set of ligatable probes (without their complementary set). They do not suggest, and apparently did not appreciate, the advantages to be obtained by combining an initial linear LCR with a subsequent full blown LCR.

Stated generally, we have discovered that the LCR can

be improved by adding a first stage comprising a controlled template-dependent reaction in which the population of templates capable of supporting template-dependent ligation in LCR is increased while substantially avoiding template-independent (i.e. in the absence of target nucleic acid) creation of detrimental amounts of template molecules capable of supporting LCR. By "detrimental amounts" we mean amounts of spurious templates which limit sensitivity in comparison to LCR without the initial controlled reaction.

In one branch of the invention, the controlled reaction comprises template-dependent ligation controlled or weighted to substantially avoid detrimental blunt end ligation.

Specifically, in one preferred method of controlled ligation, the initial template-dependent ligation is heavily weighted to selectively produce ligation products having a sequence of or corresponding to (e.g., identical to or made by ligating probes hybridizable with) the target strand, over products having a sequence corresponding to the target complement strand. (We refer to the "two target strands" or "target and target complement" to designate not only the situation in which the original target is double stranded, but also recognizing that even where the target is initially single stranded, a ligation reaction will rapidly produce a ligation product which is hybridizable to at least part of the single-stranded target. Accordingly, we use the terms first and second target strand interchangeably with target and target complement in describing assays for targets which originally are present only as a single strand.)

In this preferred embodiment, one strand of the target is effectively the predominant template controlling ligation because effectively little or no means (i.e. reagents) are provided to use the target-generated ligation products for further amplification. In the preferred extreme case, no means is provided for such secondary ligation at all and, in the controlled reaction, ligation is a substantially linear

function of the number of ligation cycles. For that reason, we sometimes refer to the initial controlled ligation stage as linear amplification of the phosphodiester bonds or "linear preamplification".

The first reaction usually consists of 10-50 cycles, although up to 100 cycles or more may be performed.

When we say that the number of ligation events for sequences of one strand occurs "selectively" or that the reaction is "weighted" in favor of ligation events for one target strand, we mean that, in the first stage, product is produced preferentially from one target strand over the second target strand. This is done primarily by varying the ratio of reagents (probes) selectively operating on the two respective strands. The ratio may be controlled in a continuum from 1:1 (even amounts of reagents operating on each strand) to infinity (no reagents operating on one of the strands). In a preferred embodiment of the first branch of the invention, the controlled reaction features providing a set of the primary nucleic acid probes and ligating the primary probes in a template-dependent manner to create additional molecules hybridizable with the first target sequence. No "secondary" probes are provided.

Another preferred embodiment under the first branch of the invention takes advantage of the fact that DNA ligases can exhibit a decided preference for ligating nicked sites (i.e., for template-dependent ligation) over blunt-end ligation. Specifically, the desired template-dependent ligation is far more efficient, and its efficiency falls less rapidly with decreasing concentration than does the efficiency of blunt end ligation. Therefore, if the concentration of at least one set of probes is kept very low during the initial controlled stage, then the chance of blunt end ligation becomes vanishingly small. The concentration of the selected probe set may be reduced by as much as 100 fold, although about 10 fold appears to be sufficient. For example, we prefer a concentration for each probe of $10^{11}$ probes or less per standard reaction vessel

(20-50uL) during the controlled reaction stage. This compares with standard LCR probe concentrations of about $10^{12}$ per 50 uL. While the rate of increase in ligated probes is reduced by reducing probe concentration, such a reduced rate can be tolerated in the initial controlled stage, which is performed to increase the number of templates initially present for standard LCR, rather than to provide directly detectable amounts of ligated species. In sum, we have taken advantage of the difference in relative efficiency of undesired blunt end ligation versus the desired template dependent ligation to produce LCR templates in a target-dependent manner, with a relatively minimal risk of blunt end ligation, thereby improving the result of the subsequent LCR reaction.

The two preferred embodiments of the first branch of the invention represent extremes of a continuum--i.e., standard concentrations of probes for one target strand with no complementary probes, on the one hand, and low concentrations of probes for both strands on the other hand. Those skilled in the art will appreciate that the invention includes points intermediate on this continuum i.e., using probes for both strands, but controlling concentration so that the population of probes for one strand is favored, and the likelihood of probe duplex formation and blunt end ligation is not detrimental. For example, fewer than $10^{11}$ molecules of probe per 50 uL vessel of at least one of the sets of probes may be present as described above.

A second branch of the invention features an initial controlled stage employing nucleic acid (preferably DNA) polymerization to generate the templates for LCR. Since polymerization is template dependent, creation of polymerization extension products generally will require the presence of the intended target sequence. To the extent that polymerization extension products are created from sites other than the intended target, such products will not significantly affect the subsequent LCR, because they will not serve as

ligation templates. Accordingly, standard PCR, or polymerization weighted to selectively amplify one strand over the other, can serve as the initial controlled reaction. A significant advantage of using PCR as an initial controlled reaction to be followed by LCR is the resulting increased specificity over that achieved with PCR alone, by filtering out (eliminating) signal from spurious PCR extension products.

In one embodiment of the second branch of the invention, the controlled reaction comprises providing copies of a single primer nucleic acid sequence complementary to a portion of one strand of the target molecule (TM) 3' to the ligation point that will occur in subsequent LCR, together with a nucleic acid polymerase and a supply of nucleoside triphosphates (by which we mean to include and even prefer deoxribonucleoside triphosphates) suitable for template-dependent polymerization. Contrary to PCR, little, and preferably none, of the primer for the other strand customarily used in PCR is provided. This process has been described as "asymmetric PCR".

It will be recognized that this weighting of the polymerase-based controlled stage is a continuum just as it was for the ligase-based controlled stage. Those reagents are used in repeated cycles of first hybridizing the primer to the target nucleic acid sequence and then reacting the hybrid with the polymerase and nucleotide triphosphates to yield an extension product/template. Denaturation yields a single stranded extension product sequence hybridizable with a sequence of the first target strand. The cycle is then repeated a number of times.

The two controlled reactions described above (polymerization and ligation) can be performed serially. Preferably polymerization is performed first.

Also, when using either type of controlled reaction (polymerization or ligation), the reaction can be weighted to produce sequences corresponding to one strand as described

above. A second parallel controlled reaction may be conducted in a separate vessel, the second reaction being weighted to produce sequences corresponding to the second target strand. The products of the two controlled amplification stages are then mixed so that the population of sequences of the two target strands subjected to LCR are approximately equal.

After any one of the controlled reaction embodiments, a standard LCR assay is used to achieve geometric amplification, in which probes with sequences corresponding to both target strand sequences are ligated, and each ligation product serves as a template for further ligation. Specifically, in the geometric stage, primary nucleic acid probes hybridize to adjacent portions of one target nucleic acid sequence and are ligated in a reaction that depends on the presence of a template (either the complement of the original target or ligation products created by ligation of probes in a manner dependent originally on the presence of the original target sequence as a template). Next, secondary nucleic acid probes hybridized to the ligated primary probes are ligated in a template-dependent manner to yield additional templates for ligation of the primary probes in a later cycle. Thus, the ligation products of the first amplification stage are submitted to geometric ligation through LCR.

A third aspect of the invention features a linear preamplification/LCR kit for performing the alternate (ligation-based) linear preamplification described above, comprising a DNA ligase, at least one of the sets of nucleic probes, a template-dependent polymerase, the above-described primer sequence, and a supply of nucleoside triphosphates suitable for polymerization by the polymerase. Preferably both the ligase and the polymerase are thermostable.

The invention permits a substantial increase in sensitivity by increasing the number of molecules containing a target sequence in a step that avoids the creation of spurious (target-independent) signal from blunt end ligation. The

invention also provides the ability to control reagent concentrations and other preamplification variables such as cycle number to optimize final LCR signal sensitivity.

One way to evaluate the benefit of the invention is as follows. For the reasons given above, a spurious background signal almost invariably will develop in LCR after a sufficient number of cycles, even in the absence of the target sequence. The assay can be evaluated by the "window" (in terms of number of cycles) bounded by the number of LCR cycles that will reliably develop a signal in the presence of target and the number of cycles likely to generate a signal in the absence of target. According to the invention, through controlled template dependent reaction prior to LCR, the "window" is enlarged by reducing the number of LCR cycles necessary to develop a signal without substantially reducing the number of cycles at which a background signal appears in the absence of target.

Alternatively, the improvement of the invention may be evaluated in terms of improved sensitivity of an assay. In other words, compared to standard LCR, initial controlled amplification according to the invention improves the ability to distinguish samples containing target from samples without target.

Other features and advantages of the invention will be apparent to those skilled in the art from the following descriptions of the preferred embodiment.

Brief Description of the Drawings

Fig. 1 is a diagram representing the linear preamplification steps in one embodiment of the invention.

Fig. 2 is a diagram of LCR following the linear preamplification of either Fig. 1 or Fig. 3.

Fig. 3 is a diagram representing the linear preamplification steps in a second embodiment of the invention.

Fig. 4 is a graph of the results of assays described in

Example 4.

Fig. 5 is a graph of the results of assays described in Example 5.

In Fig. 1, a single-stranded nucleic acid target molecule (e.g., DNA) ("TM") contains a known nucleic acid sequence target sequence ("TS"). (Note that TS need not be all of TM). The molecule TM may occur in the sample as a single stranded molecule; alternatively, the sample may contain a double stranded (hybridized) target molecule, in which case we arbitrarily denote one strand TM and the other TM'. Similarly TS' designates the sequence complementary to TS.

Controlled amplification according to the invention is depicted in Fig. 1 by target-dependent ligation of probes P1 and P2 hybridized to TS (step A, Fig. 1). By denaturation (step B, Fig. 1), the ligated entity P1.P2 is separated from TM. Note that this entity (P1.P2) cannot serve as a template for ligation of P1 and P2.

Each cycle comprises:

a) providing single stranded TS;

b) hybridizing P1 and P2 to TS;

c) ligating P1 to P2 in a template(TS)-dependent ligation reaction to yield a TS-P1.P2 duplex

d) denaturing the resulting TS-P1.P2 duplex yielding the original target molecule with sequence TS and the ligated complementary molecule P1.P2. In each cycle, there will be a particular yield of P1.P2, and that yield will be a function of the presence original target TS only, because no TS sequences are created during the linear phase. P1 and P2 effectively do not contribute to spurious background through blunt end (templated-independent) ligation because of the absence of complementary P1' and P2' sequences.

After the desired number of linear preamplification cycles, the LCR step (Fig. 2) involves addition of probes P1' and P2', in order to amplify the P1.P2 product of linear

preamplification. Geometric LCR amplification in the presence of the P1.P2 product will yield a signal within a predictable number of cycles. In many situations, when LCR is applied directly to the target without initial controlled amplification, a greater number of LCR cycles would be required to yield that signal. In that way the invention substantially reduces the chance that undesired blunt end ligation products will be present in sufficient amount to yield an LCR signal in the cycle range producing a target-generated signal.

In Fig. 3 linear preamplification is achieved by polymerization, in contrast to the ligation shown in Fig. 1. Again, the target sequence TS of target molecule TM is used to produce additional copies of the complementary sequence TS'. A primer sequence Pr is supplied which includes a sequence C complementary to the 3' end of TS. Of course, the primer need only hybridize to the target molecule (TM) 3' to the ligation point that will occur in subsequent LCR. Sequence C is hybridized to TS, and polymerase and nucleotide triphosphates are added to extend sequence C in a template-dependent fashion, creating Pr.TS' containing an additional copy of TS', complementary to TS. As in the controlled ligation-based reaction, the result of multiple cycles is multiple copies of TS', but those copies do not themselves serve as template for further amplification during the linear stage, since a second primer is omitted.

Controlled preamplification techniques employing either ligase or polymerase enzymes are useful to increase the sensitivity of LCR. For example, controlled initial amplification increases the signal/noise ratio compared to a control that does not include controlled initial amplification. When evaluated in terms of improved sensitivity of an assay compared to a standard LCR assay, controlled amplification may improve the sensitivity by 10 fold or more thus enhancing the ability to distinguish samples containing the desired target from samples without the target.

In general, the controlled amplification is run for 10-100 cycles. The sample may then be heat treated at a temperature sufficient to destroy the activity of the enzyme present in the sample. If the initial controlled amplification is a ligased-based amplification, this step is optional. If necessary, the reaction mixture is then suitably adjusted (buffer, pH, etc.) in a manner known to those skilled in the art to support the LCR reaction.

Additional oligonucleotides are added to a concentration sufficient to serve as a complete LCR probe set, fresh ligase may be added, and LCR is performed on the target enriched sample. In general, oligonucleotides are used in the geometric phase in equimolar amounts at a final concentration of 10-100nM, more usually 30-100nM. Further considerations regarding oligonucleotides, reagents, and cycling conditions for the ligase linear amplification reaction are the same as those used for geometric LCR and are generally disclosed in EP-A-320 308. Conditions which can be used for controlled polymerase reactions are generally disclosed in Maniatis et al. Molecular Cloning (1982 Cold Spring Harbor); and Panet and Khorana (1974) J. Biol. Chem. **249**:5213-5221. This polymerization differs from that described in U.S. patent Nos 4,683,202 and 4,683,195, in incorporated herein by reference, in that only one primer is used.

The following examples are provided to illustrate the invention, not to limit it. Various improvements and modifications may be practiced with the invention.

**Example 1: Controlled Amplification By Ligation**

The following duplex target DNA sequence is present in the commercially available vector PUC19. It is presented as a single strand for simplicity sake. The "-" in the sequence represents the intended point of ligation of the probes.

3'-...TTAAGCTCGA GCCATGGG-CC CCTAGGAGAT CTCAGCTGGA CGT...-5'

The following probes set was designed to hybridize to the

target sequence for use in preamplification of the target sequence:

| | | |
|---|---|---|
| A | 5'-AATTCGAGCT CGGTACCC | ID No 1 |
| B | 5'-GGGGATCCTC TAGAGTCGACC TGCA | ID No 2 |

Primers A & B and their complements (A'&B') were prepared by standard methods using Applied Biosystems Model 380B. Primers A' and B were treated with polynucleotide kinase and ATP to render their 5' ends phosphorylated by the general method of Berger and Kimmel (1987) Guide to Molecular Cloning Techniques pp. 438 et seq. Primer A' was radioactively labeled at its 3' end by treatment with terminal transferase and $-^{32}$P-dCoTP (Cordecypin) by the general method of Tu and Cohen (1980) Gene 10:177-183.

Thermostable ligase was prepared according to the procedure described in EP-A-O 373 962. Samples were prepared which contained "1X LCR Buffer"--i.e., 50mM EPPS, pH 7.8, 100mM KCl, 10.0mM $MgCl_2$, 1.0mM dithiothreitol, 10mM $NH_4Cl$, 10ug/ml Bovine Serum Albumin--and 100uM NAD, 20ug/ml carrier DNA such as calf thymus DNA plus equimolar amounts of primers A and B (85nM each). Controlled amplification was performed on a sample containing 1000 target molecules in the presence of thermostable ligase. A single cycle encompasses the following steps and is the same for either linear or geometric ligase-mediated amplification.

a) Heat to 90°C for 1 minute to denature the DNA

b) incubate at 50°C for 1 minute to allow annealing and ligation of adjacent probes.

After 50 cycles, the reaction was stopped and ligase was denatured by boiling for 10 min. To the "preamplified" sample were added the other two probes (A' & B'; 80nM each) of the LCR probe set and new ligase, and cycling was resumed for an additional 20-50 cycles. Samples were withdrawn at various times and analyzed by polyacrylamide gel electrophoresis and

autoradiography. As controls, standard LCR reactions were performed using the four probes and either zero or 1000 targets. Samples were withdrawn and analyzed as for the pre-amplified reaction.

The control sample without target sequence produces signal with characteristic kinetics, and the sample with 1000 target sequences (but no controlled amplification) produces signal which has very similar kinetics. The signal in the target-containing sample subjected to initial controlled amplification, however, appeared at least 4 to 6 cycles earlier in the LCR (geometric amplification) procedure than the others. Thus, by beginning with controlled amplification, 1000 targets can be readily distinguished from background in a case where such distinction is difficult without preamplification. This allows identification of numbers of targets not reliably identifiable without linear pre-amplification, and represents an improved sensitivity.

**Example 2: Controlled Amplification by Polymerization**

Linear preamplification by polymerization also improves the sensitivity of the LCR assay. Accordingly, Example 2 features the use of the target and probe sequences given in Example 1. To a sample containing 1000 targets, a single oligonucleotide probe, A, is added. Thermostable DNA polymerase, buffer, and dNTPs are added according to the method described in Maniatis et al., cited above; see also the other references cited above regarding polymerization. Fifty hybridization, extension, and denaturation cycles are performed; the reaction is then stopped and the polymerase activity is destroyed by boiling for 10 minutes.

The entire preamplified sample is ethanol precipitated and centrifuged, residual salts removed by ethanol wash and the precipitate dried. The precipitate is then resuspended in the standard LCR reaction mix, probes A, B, A', B' (80nM for each probe) are added, thermostable ligase is added, and LCR is

performed. Controls containing 1000 or zero targets are also run.

Example 3: Serial Controlled Amplification

Using the target and probe sequences given in Example 1, a single oligonucleotide probe, A, is added to a sample containing 1000 targets. Polymerase-mediated preamplification is performed as in Example 2 including precipitation, washing, resuspension in LCR buffer and addition of ligase enzyme. As opposed to Example 2, only probes A' and B' are added, probe B is omitted. In the absence of probe B, geometric amplification cannot occur and thus a second linear amplification step is initiated. The linear reaction is performed for 50 cycles at which point the reaction is stopped as in Example 1, and fresh thermostable ligase and probe B are added and the geometric amplification is performed for 20-50 additional cycles. Controls may be run as in the previous Examples.

Example 4: Controlled Amplification with Detection on the IMx® instrument

Duplex DNA having the following sequence (adapted from the L1 region of Human Papilloma Virus Type 16) was used as a target to study the linear preamplification process. The sequence is presented as a single strand for simplicity with a hyphen designating the ligation site:

5'-AGGTTGTAAG CACGGATGAA TATGT-TGCAC GCACAAACAT ATATTATCAT G-3'

The following probe set was designed to hybridize to the above target sequence for use in preamplifiction of the target sequence. The probes were synthesized by standard techniques (Fl=fluorescein and Bio=biotin).

| | | |
|---|---|---|
| 179.1 | Fl-AAGTTGTAAG CACGGATGAA TATGT-3' | ID No 3 |
| 179.2 | 5'-ACATATTCAT CCGTGCTTAC AACT-Fl' | ID No 4 |
| 179.3 | 5'-TGCACGCACA AACATATATA TTACA-Bio | ID No 5 |
| 179.4 | Bio-ATGATAATAT ATGTTTGTGC GTGCA-3' | ID No 6 |

The 5' end of probe 179.1 was haptenated with fluorescein

for use in the IMx MEIA assay as follows. A solution of 15 uL of 0.9 $A_{260}$ Unit/uL of oligo 5' XAAGTTGTAAGCACGGATGAATATGT-3' (ID No 3), where X is Aminomodifier II(™) (Clontech), was treated with 4 mg of fluorescein isothiocyanate (FITC, Kodak) dissolved in 485 uL of sodium borate buffer, pH 9.2, at room temperature for 15h in the dark. The excess FITC was removed on a NAP-5 column (Pharmacia), and the eluate was concentrated to a volume of 100 uL on a Speedvac(™) (Savant Instruments). The solution was diluted to 200 uL with formamide, and the sample was separated by electrophoresis at 40 W constant power on a 1.5 mm thick, 12% acrylamide/8M urea gel. The electrophoresis was stopped after 5h, and the product band on the gel was visualized with long-wave UV shadowing. The product had a lower mobility than unlabeled starting material. The excised FITC-oligo band was extracted overnight with 3 ml of 1.0 M triethylammonium acetate, and the aqueous extracts were lyophilized. Reconstitution of the residue in distilled water was followed by NAP-5 desalting. After haptenation, the probe was diluted to 1012 molecules per uL in Tris-EDTA buffer (10mM Tris, 1mM EDTA).

The 3' end of probe 179.2 (ID No 4) was haptenated with fluorescein as described above, using 3'-amino CPG (Glen Research) in place of 5' Aminomodifier II(™). The probe is treated with polynucleotide kinase and diluted to $10^{12}$ molecules per uL.

The 3' end of probe 179.3 was haptenated as follows: A solution of 35uL of 0.378 $A_{260}$ Unit/uL of oligo 5'-TGCACGCACAAACATATATTATCAX-3', (ID No 5) where X is 3'-amino CPG (CPG= controlled pore glass; Glen Research), was treated with 10 mg of Biotin-(aminocaproyl)2-NHS active ester (Clontech) in 215/250 uL 100 mM pH 7.5 phosphate/DMF at room temperature for 15h. Workup and electrophoresis as in the fluorescein case gave a product band on short wave UV shadowing which migrated the equivalent of 4 bases slower than the starting material. Excision and extraction, followed by desalting as in the case

of fluorescein. The probe was treated with polynucleotide kinase and ATP to phosphorylate the 5' end, and it was diluted to $10^{12}$ molecules per uL.

The 5' end of probe 179.4 (ID No 6) was haptenated as described above and diluted to $10^{12}$ molecules per uL.

Each linear preamplification was run in a 50uL volume as follows:

| | (uL) |
|---|---|
| dH2O | 32.75 |
| 5X LCR Buffer | 10.0 |
| 10mM NAD | 0.5 |
| 179.1 ($10^{12}$ molecules per uL) | 0.5 |
| 179.3 ($10^{12}$ molecules per uL) | 0.75 |
| target (varied concentrations) | 1.5 |
| 1X ligase (from 2 to $10x10^6$ units where 1 unit seals 1uM nicked DNA per minute) | 1.0 |
| | 50.0 |

Where possible, reaction volumes were created from larger, pooled volumes that were divided into aliquots. Each individual reaction was overlayed with 30uL of mineral oil.

Each reaction aliquot (minus the ligase) was subjected to 100°C for 3 minutes to denature the target, followed by 85°C for 30 sec and 50°C for 20 sec. Then 1uL of ligase was added. The preamplification reaction was run for 50 thermal cycles (30 secs at 85°C followed by 20 secs at 50°C).

The LCR geometric amplification phase was accomplished by adding a 3uL aliquot of a pool consisting of:

| | |
|---|---|
| $dH_2O$ | 1.75uL |
| 179.2 ($10^{12}$ molecules/uL) | 0.75uL |
| 179.4 ($10^{12}$ molecules/uL) | 0.5uL |

LCR reactions were run for 30 cycles. Appropriate controls were run.

A comparison of LCR under comparable conditions with

and without linear preamplification showed that linear preamplification significantly increased sensitivity as shown in Fig. 4.

**Example 5: Reduced Initial Probe Concentration**

The following example illustrates template-dependent ligation using reduced probe concentration as a method for controlling blunt end ligation.

The following two stock solutions were prepared.

STOCK SOLUTION #1

6.3 uL 10x LCR buffer
4.5 uL 10 mM NAD
6.2 uL $H_2O$
4.5 uL each probe at $10^{11}$ oligonucleotides/uL
10.0 uL $CoT^{32}P$-(Cordecypin) labelled probe.

STOCK SOLUTION #2

2.0 uL 10x LCR buffer
4.5 uL each probe at $10^{12}$ oligonucleotides/uL.

An aliquot of 10 uL of SOLUTION #1 was added to each of four(4) 0.65 ml Eppendorf tubes together with 3.0 uL of background DNA with or without target to each tube. LCR reaction was initiated and continued for 20 cycles. The tube was spun briefly in a centrifuge and returned to 90°C. 4.0 uL of SOLUTION #2 was added to each tube and temperature cycling was continued as with standard LCR.

Fig. 5 depicts the results of the control (standard LCR) and of the use of an initial controlled (low probe concentration) ligation ("ICL").

**Example 6: PCR as the Initial Controlled Reaction**

A segment of the cystic fibrosis (CF) gene was amplified by PCR from genomic DNA samples obtained from several CF families using primers A and B (Table 1). PCR was performed for 30 cycles. Each cycle consisted of a 60 sec incubation at 94°C, a 43 sec incubation at 62°C, and a 120 sec incubation at 72°C.

Following PCR, reaction products were diluted 1:200 with water and 1uL aliquots were used as target for LCR reactions with probe sets specific for the normal CF allele and the major defective CF allele which contains a three nucleotide deletion (Riordan, J.R. et al. Science 245:1066 (1989); Kerem, B. et al. Science 245:1073 1989). LCR probes used for amplification of the normal and deleted CF alleles are listed in Table 1. Probes C', E', and D' were phosphorylated as in Example 1, above. LCR was performed for 25 cycles. Each cycle consisted of a 30 sec incubation at 85°C and a 20 sec incubation at 50°C. For identification of the normal allele, oligonucleotides C' and D were used at $7.5 \times 10^{11}$ molecules/reaction and oligonucleotides E and D' at $5 \times 10^{11}$ molecules/reaction. For the deleted allele, oligonucleotides C' and D were used at $7.5 \times 10^{11}$ molecules/reaction and oligonucleotides E and D' at $5 \times 10^{11}$ molecules/reaction. Reaction volume in each case was 50uL. Following LCR, reaction products were analyzed by the IMx® MEIA assay. Selected representative results are shown in Table 2. Although all patients were "PCR positive", LCR distinguished those who carried the CF allele from those who did not.

This Example demonstrates the improvement of the invention of the invention if the normal allele is viewed as "true" target and the deleted allele is viewed as spurious extension.

Table 1

| | | |
|---|---|---|
| PCR Primers: | | |
| A: | 5'-GTTTTCCTGG ATTATGCCTG GGCAC-3' | ID No 7 |
| B: | 5'-GTTGGCATGC TTTGATGACG CTTC-3' | ID No 8 |
| LCR Probes: | | |
| Normal Allele | | |
| C: | Fl-CACCATTAAA GAAAATATCA TCTT-3' | ID No 9 |
| C': | 5'-AAGATGATAT TTTCTTTAAT GGTGC-Fl | ID No 10 |
| D: | 5'-TGGTGTTTCC TATGATGAAT ATAGA-Bio | ID No 11 |
| D': | Bio-CTATATTCAT CATAGGAAAC ACCA-3' | ID No 12 |

Deleted Allele

| | | |
|---|---|---|
| E: | Fl-TGGCACCATT AAAGAAAATA TCAT-3' | ID No 13 |
| E': | 5'-ATGATATTTT CTTTAATGGT GCCAG-Fl | ID No 14 |
| D: | 5'-TGGTGTTTCC TATGATGAAT ATAGA-Bio | ID No 15 |
| D': | Bio-CTATATTCAT CATACGAAAC ACCA-3' | ID No 16 |

Table 2

| | IMx Signal (counts/sec/sec) | |
|---|---|---|
| | Normal | Deleted |
| No target | 19 | 97 |
| patient 1 | 1826 | 107 |
| patient 2 | 1973 | 1997 |
| patient 3 | 24 | 1946 |

Other embodiments are within the following claims. For example, the low probe concentration controlled amplification of Example 5 may be combined with one or more of the weighted amplification stages described above.

What is claimed is:

1. A method for performing a ligase chain reaction (LCR) on target nucleic acid in a sample, wherein a first template-dependent controlled reaction, prior to said LCR, increases the population of molecules capable of participating as templates in said LCR selectively in the presence of said target nucleic acid, while substantially avoiding creation of detrimental amounts of template molecules capable of participating in said LCR in the absence of said target nucleic acid, characterized in that said template-dependent controlled reaction comprises:

providing a first set of ligatable nucleic acid probes capable of hybridizing to a target strand, in substantial excess over a second set of ligatable probes (if any) capable of hybridizing to complement of the target strand or to ligated first probes; and ligating said members of said first set of probes in a template-dependent manner, the second set of probes (if any) being provided in a concentration low enough to substantially avoid detrimental ligation independent of target.

2. The method of claim 1 in which substantially no amplification products containing a sequence of the target complement strand are generated in said controlled reaction, and substantially the only template for said first controlled reaction is said target nucleic acid.

3. The method of claim 2 further comprising a parallel template-dependent controlled reaction performed in a separate vessel from the vessel used for said first controlled reaction, said parallel controlled reaction comprising a reaction weighted to generate products containing a sequence corresponding to said target complement strand selectively with substantially no products containing a sequence corresponding to the target strand and said method further comprising mixing products of said first controlled reaction with products of said parallel controlled reaction, and performing LCR on said mixture.

4. The method of claim 1 in which the template-dependent controlled reaction comprises a ligase chain reaction using a concentration for at least one of the sets of LCR probes low enough to substantially eliminate blunt end ligation.

5. The method of claim 4 wherein the concentration of one set of probes exceeds the concentration of the other set of probes by about 10 fold.

6. The method of claim 4 in which the reaction vessel has a volume of 20 to 50 uL and the concentration of both probes for said one set of LCR probes is no more than $10^{11}$ molecules.

7. The method of claim 1 in which the controlled reaction comprises a template-dependent polymerization stage and a template-dependent ligation stage, said ligation stage comprising template-dependent ligation substantially without creation of detrimental amounts of blunt end ligation products.

8. The method of claim 7 in which said polymerization stage is performed first and said ligation stage is performed second.

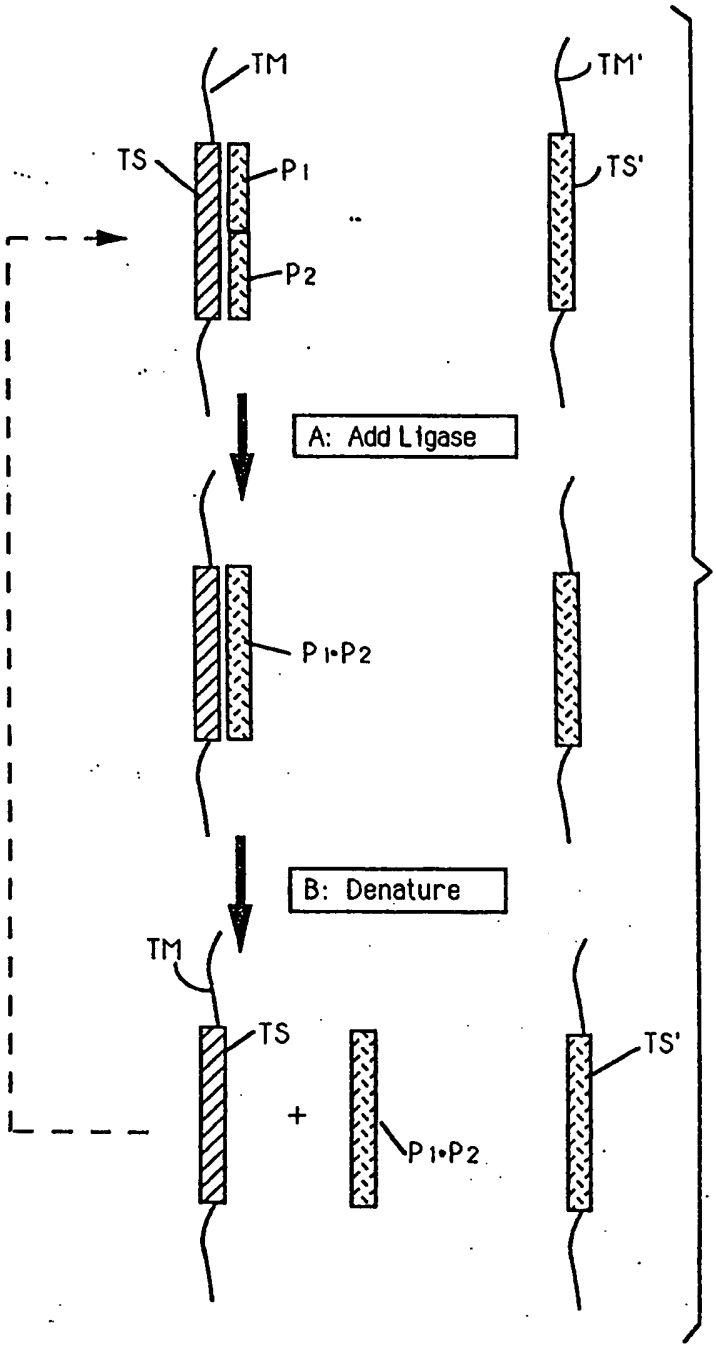


Fig. 1

TS = target seq
TM = target mol (ss)

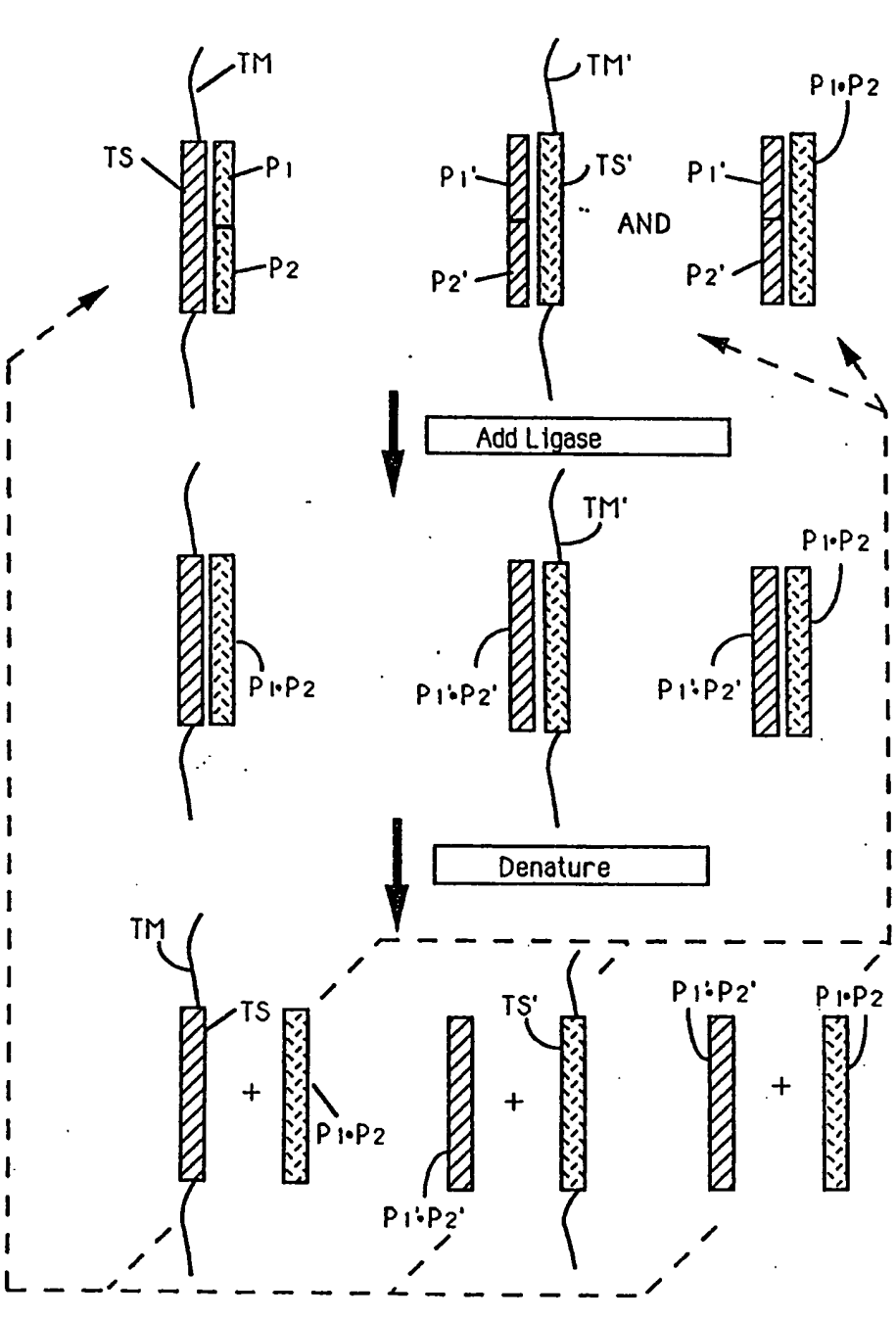


Fig. 2

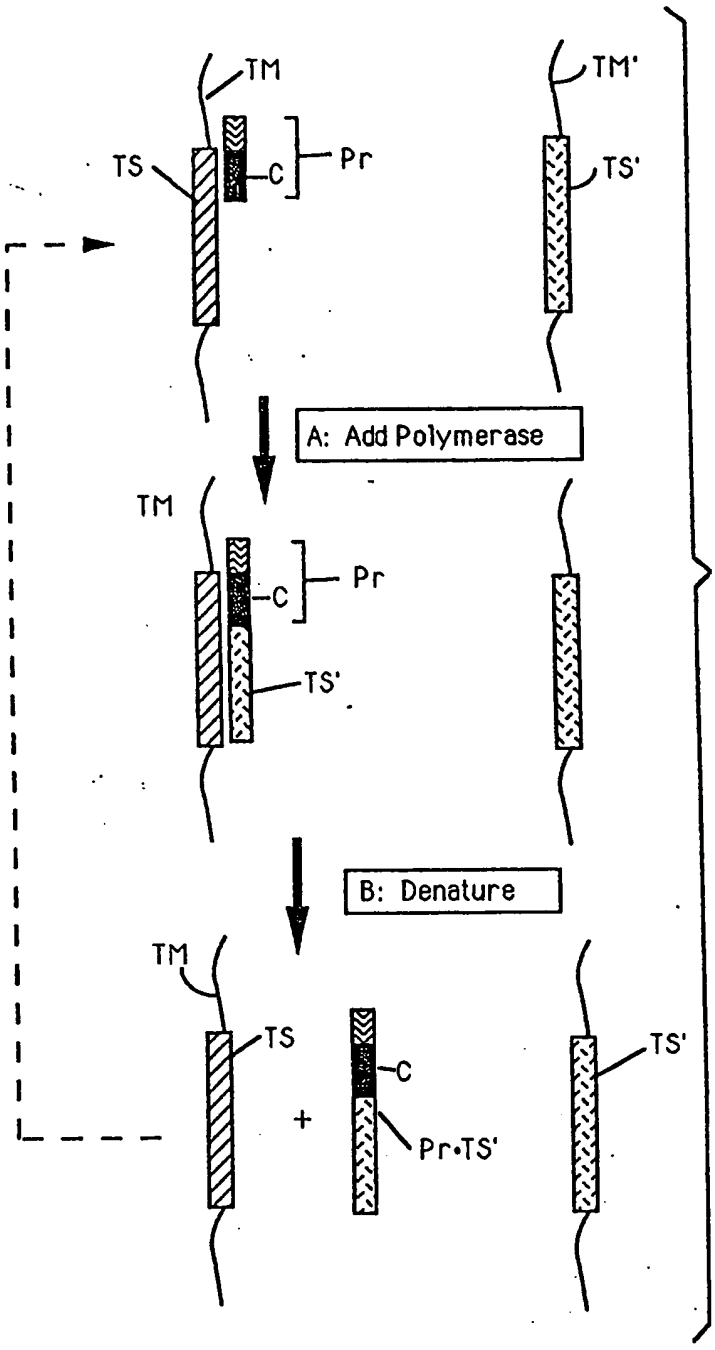
TM
TS
C
Pr
TM'
TS'
A: Add Polymerase
TM
C
Pr
TS'
B: Denature
TM
TS
C
+
Pr•TS'
TS'

Fig. 3

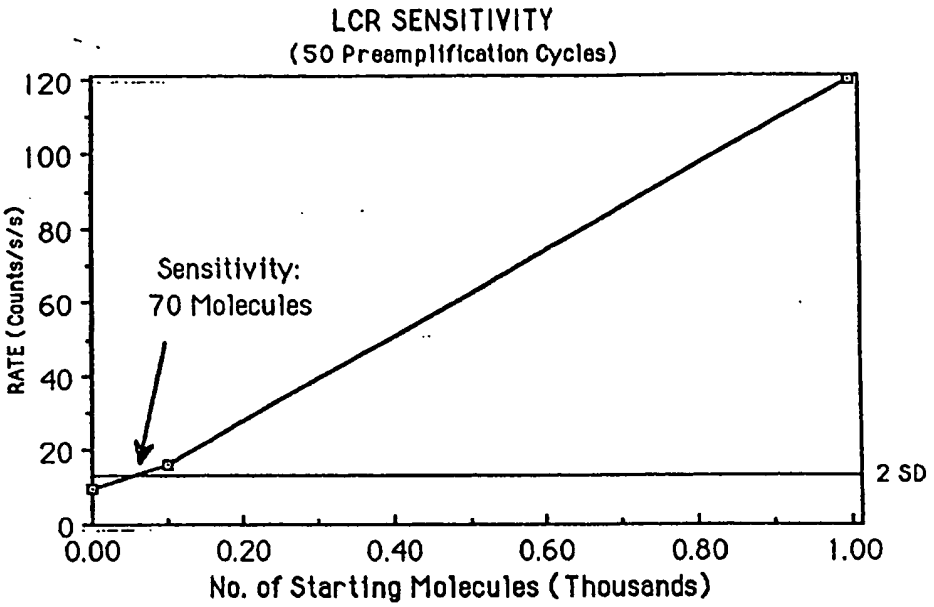
LCR SENSITIVITY
(50 Preamplification Cycles)
120
100
80
60
40
20
0
RATE (Counts/s/s)
Sensitivity:
70 Molecules
2 SD
0.00
0.20
0.40
0.60
0.80
1.00
No. of Starting Molecules (Thousands)

FIG. 4

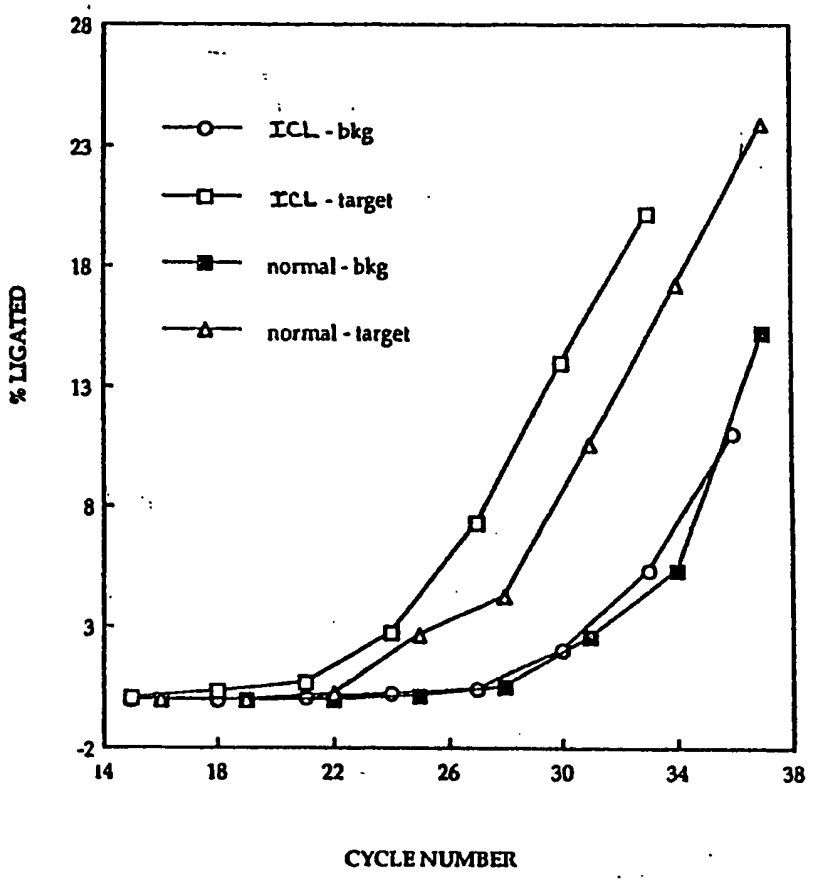
28
23
18
13
8
3
-2
% LIGATED
ICL - bkg
ICL - target
normal - bkg
normal - target
14
18
22
26
30
34
38
CYCLE NUMBER

FIG. 5